Most published and unpublished dissertations should be excluded from meta-analyses

Coyne, J.C.; Hagedoorn, M.; Thombs, B.

Published in:
Psycho-oncology

DOI:
10.1002/pon.1788

Psycho-Oncology
Psycho-Oncology **20**: 224–225 (2011)
Published online in Online Library (wileyonlinelibrary.com). **DOI**: 10.1002/pon.1788

Letter to the Editor

Most published and unpublished dissertations should be excluded from meta-analyses: comment on Moyer *et al.*

Moyer *et al.* [1] systematically collected published and unpublished dissertations evaluating psychosocial interventions for cancer patients and examined the methodological quality of these studies. They concluded that because published and unpublished dissertations differ little in methodological adequacy, the inclusion of unpublished dissertations in meta-analyses is desirable in order to avoid publication bias.

Instead of focusing explicitly on quality, however, the Moyer *et al.*'s [1] analysis combined criteria intended to reflect adequate reporting of results from trials with criteria that indeed reflect quality, or the likelihood that trial results reflect underlying clinical realities. For instance, their first four criteria concern whether the dissertation reported the number of participants approached for consent to participate in a study; reported the number initially participating; reported comparisons of participants to patients approached, but not participating; and reported the number dropping out of treatment. A dissertation could have earned a perfect score on these ratings for reporting this information, despite reporting extremely low uptake and low retention of participants, which would suggest a high probability of bias and potentially poor external and internal validity.

Nonetheless, the evidence reported by Moyer *et al.* [1] makes it clear that both published and unpublished dissertations are generally of poor quality. Dissertations with at least 10 patients per cell were included in their analyses, with mean numbers of initial participants per cell of 37.7 (SD = 42.5) in published dissertations and 29.4 (SD = 23.94) in unpublished dissertations. Moyer *et al.* do not indicate how many small, grossly underpowered studies were included in either the published or unpublished dissertations. However, the mean cell sizes and large standard deviations for mean participants per cell suggest a substantial number.

The problems posed by studies with small cell size are not widely appreciated [2]. As demonstrated by Kraemer *et al.* [3], the inclusion of small, underpowered trials in meta-analyses results in substantially *overestimated* effect estimates due to confirmatory publication bias. Statistical correction is impossible with a proportionately large number of underpowered studies. To achieve 80% statistical power to detect a moderate effect size (e.g., $\delta = 0.50$), 64 patients would need to be randomized to each of the intervention and control groups. A small study of 20 patients per group would have only 34% power to detect a moderate effect size. With 20 patients per group, a fairly large effect size of 0.65 would be needed just for statistical significance. The problem is even worse than that; however, as small studies with true null effects that cross the $p < 0.05$ threshold do it by varying degrees. With 20 patients per group and a true null effect, the expected standardized effect size in a meta-analysis of statistically significant trials would be 0.90–1.00. Thus, albeit counter-intuitive, grossly underpowered studies with positive results, including most published and unpublished dissertations, are most often false positives.

Cuijpers *et al.* [4] recently showed that, when only high-quality studies were considered, the effect estimates for psychotherapy for depression decreased from large ($d = 0.74$) to small ($d = 0.22$). Quality criteria included sample size, use of intention-to-treat analyses, independent randomization, utilization of treatment manuals, and treatment integrity. Of the studies reviewed by Moyer *et al.*, 17% of published dissertations and 38% of non-published dissertations were not randomized trials at all; only 12 and 6% of published and unpublished dissertations, respectively, used intent-to-treat analyses; only 11 and 18% described a specific method of randomization and measures to prevent subterfuge; fewer than half in either group used treatment manuals; and only 67 and 49% monitored intervention implementation. Most studies reported that they assessed baseline equivalence, but Moyer *et al.* do not report whether or not this was achieved. Indeed, findings of baseline equivalence of intervention and control groups, based on the lack of statistically significant differences, are often meaningless with small studies because there is too little power to detect differences that may be individually or collectively decisive in determining the outcome of a trial.

The literature concerning psychosocial interventions for cancer patients has been shown to have serious shortcomings in terms of methodology [5] and clinical and statistical heterogeneity [6]. The pervasiveness of these problems raises concerns about whether studies should automatically be included in meta-analyses based simply on their availability [2] or even whether a summary estimate

of effect size derived from the overall literature is meaningful [7]. Statistical adjustment or other methods to assess the influence of poor-quality studies on the results of meta-analyses do not work when many or most studies share similar methodological problems and are of generally low quality. Solutions to the problem of adequately gauging the efficacy of psychosocial studies will not be found by introducing even more small, methodologically flawed studies into consideration. Instead, we need to rely on methodologically stronger studies with adequate sample sizes.

The uncritical inclusion of unpublished dissertation studies in meta-analyses should be discouraged. A more judicious decision would be to base inclusion in meta-analyses on study quality or, at a minimum, to present results for high- and low-quality studies separately [8]. Based on the results presented by Moyer *et al.*, most dissertations, regardless of their publication status, would be graded as low quality based on any of the commonly used quality assessment tools.

References

1. Moyer A, Schneider S, Knapp-Oliver SK, Sohl SJ. Published versus unpublished dissertations in psycho-oncology intervention research. *Psycho-Oncology* 2010; **19**:313–317.
2. Coyne JC, Thombs BD, Hagedoorn M. Ain't necessarily so: review and critique of recent meta-analyses of behavioral medicine interventions in health psychology. *Health Psychol* 2010;**29**:107–116.
3. Kraemer HC, Gardner C, Brooks JO, Yesavage JA. Advantages of excluding underpowered studies in meta-analysis: inclusionist versus exclusionist viewpoints. *Psychol Methods* 1998;**3**:23–31.
4. Cuijpers P, van Straten A, Bohlmeijer E *et al.* The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. *Psychol Med* 2009;**40**:211–223.
5. Coyne JC, Thombs BD, Hagedoorn M. A meta-analysis of psychosocial interventions for cancer patients gone awry. *Ann Behav Med* 2009;**37**:94–96.
6. Coyne JC, Lepore SJ, Palmer SC. Efficacy of psychosocial interventions in cancer care: evidence is weaker than it first looks. *Ann Behav Med* 2006;**32**:104–110.
7. Newell SA, Sanson-Fisher RW, Savolainen NJ. Systematic review of psychological therapies for cancer patients: overview and recommendations for future research. *J Natl Cancer Inst* 2002;**94**:558–584.
8. Tak LM, Meijer A, Manoharan A, de Jonge P, Rosmalen JG. More than the sum of its parts: meta-analysis and its potential to discover sources of heterogeneity in psychosomatic medicine. *Psychosom Med* 2010;**72**:253–265.

James C. Coyne[1], Mariet Hagedoorn[2] and Brett Thombs[3]

[1]*School of Medicine, University of Pennsylvania, Psychiatry, 3535 Market St, Room 376, Philadelphia, PA, USA*

E-mail: jcoyne@mail.med.upenn.edu

[2]*Department of Health Sciences, Section Health Psychology, University Medical Center Groningen, Groningen, The Netherlands*

E-mail: Mariet.Hagedoorn@med.umcg.nl

[3]*Department of Psychiatry, McGill University, Montreal, Quebec, Canada*

E-mail: brett.thombs@mcgill.ca

DOI: 10.1002/pon.1788